Research Article

A Categorical Structure-Activity Relationship Analysis of the Developmental Toxicity of Antithyroid Drugs

Albert R. Cunningham,[1, 2, 3] C. Alex Carrasquer,[1, 3] and Donald R. Mattison[2, 3]

[1] *James Graham Brown Cancer Center, University of Louisville, 529 S. Jackson Street, Louisville, KY 40202, USA*
[2] *Departments of Medicine and Pharmacology and Toxicology, University of Louisville, Louisville, KY 40202, USA*
[3] *Eunice Kennedy Shriver National Institute of Child Health and Human Development, Bethesda, MD 20892, USA*

Correspondence should be addressed to Albert R. Cunningham, al.cunningham@louisville.edu

The choice of therapeutic strategies for hyperthyroidism during pregnancy is limited. Surgery and radioiodine are typically avoided, leaving propylthiouracil and methimazole in the US. Carbimazole, a metabolic precursor of methimazole, is available in some countries outside of the US. In the US propylthiouracil is recommended because of concern about developmental toxicity from methimazole and carbimazole. Despite this recommendation, the data on developmental toxicity of all three agents are extremely limited and insufficient to support a policy given the broad use of methimazole and carbimazole around the world. In the absence of new human or animal data we describe the development of a new structure-activity relationship (SAR) model for developmental toxicity using the cat-SAR expert system. The SAR model was developed from data for 323 compounds evaluated for human developmental toxicity with 130 categorized as developmental toxicants and 193 as nontoxicants. Model cross-validation yielded a concordance between observed and predicted results between 79% to 81%. Based on this model, propylthiouracil, methimazole, and carbimazole were observed to share some structural features relating to human developmental toxicity. Thus given the need to treat women with Graves's disease during pregnancy, new molecules with minimized risk for developmental toxicity are needed. To help meet this challenge, the cat-SAR method would be a useful in screening new drug candidates for developmental toxicity as well as for investigating their mechanism of action.

1. Introduction

Hyperthyroidism occurs in approximately 2/1000 pregnancies with the majority of those cases being Graves' disease [1], an autoimmune disorder caused by TSH-receptor stimulating autoantibodies. Propylthiouracil (PTU) is the primary treatment of hyperthyroidism in the US followed by methimazole (MMI) while carbimazole (CMI) is not distributed in the US. Surgery and radioiodine are not recommended treatment modalities during pregnancy. Untreated hyperthyroidism during pregnancy leads to developmental toxicity which includes spontaneous abortion, prematurity, growth restriction, and fetal death [1, 2]. It is not clear if untreated hyperthyroidism during pregnancy leads to structural malformations [3, 4]; however one dataset suggested that the risk of malformations was greater than expected among infants from untreated than treated

pregnancies [5]. Additionally, there is concern that one of the medical therapeutic strategies, MMI, may be a weak developmental toxicant, producing structural malformations [6], as well as fetal goiter. Part of this concern stems from the observation that CMI is a weak developmental toxicant and one of its metabolites is MMI.

On the one hand, as just mentioned, treatment of hyperthyroidism during pregnancy may result in developmental toxicity, as the available drugs cross the placenta and can cause fetal goiter. On the other hand, while the data remain unclear, there is evidence suggesting that maternal hypothyroidism is associated with impaired fetal neurodevelopment. Consequently the clinician must balance the use of the antithyroid medications against the potential developmental consequences of inadequate or aggressive therapy, with a limited set of therapeutic options.

Several assumptions about PTU and MMI have guided the medical treatment strategies for Graves' disease during pregnancy in the US [2, 6–8]. It has been assumed that placental transport of PTU is less than MMI, it has been assumed that fetal and neonatal thyroid effects of PTU is less than with MMI (based on lower placental transport), it has been assumed that MMI exposure during pregnancy leads to greater impairment of neurocognitive development than with PTU exposure, and it has been assumed that MMI produces structural defects and that PTU does not. A thoughtful review and analysis of literature by Mandel has demonstrated that these assumptions are not correct [2, 6].

Given the uncommon nature of Graves' disease in pregnancy and the weak link between the antithyroid medications and malformations, we have taken a different approach to assess the potential developmental toxicity of the drugs which are being used to treat the disease in pregnancy. Several years ago we created datasets to evaluate the structural determinants of developmental toxicity in experimental animals and humans [9–13]. Statistical analyses demonstrated that animal models are reasonable predictors of human developmental toxicity [14], and that rules could be agreed upon among experts in developmental toxicity for evaluating animal and human data [9]. Subsequently we evaluated the utility of structure-activity relationship (SAR) models generated by MultiCASE for studying and predicting developmental toxicity in diverse species including humans [12]. This later dataset of chemicals assessed for human developmental toxicity has subsequently been utilized to create a more transparent and robust model of developmental toxicity using the categorical-SAR (cat-SAR) expert system.

Briefly, the cat-SAR expert system diverges from other SAR expert systems wherein there is a high degree of user flexibility in both learning set development and model parameterization [15]. Cat-SAR analysis allows the user to specify adjustable modeling attributes including the selection of size of the 2-dimensional fragments, whether or not to include hydrogen atoms in the analysis, and rules for identifying important fragments for the final model. Hence, the selection of compounds included in the learning set and control over various model attributes provides the user with the ability to rigorously explore the relationships between chemical structure and biological activity. Application of the cat-SAR expert system to a toxicological or pharmacological endpoint is thus not constrained wherein a given set of data must fit the attributes of a predefined and often proprietary modeling process.

2. Methods

2.1. Human Developmental Toxicity Dataset.
Data on human developmental toxicity were derived from the teratogen information system and a database that utilized the US FDA guidelines as described previously [12]. The chemicals in this database were specifically characterized with respect to risk for human developmental toxicity including death, growth retardation and functional and structural abnormalities.

2.2. Cat-SAR Structure Activity Relationship Expert System.
The cat-SAR approach is a computational SAR or in silico toxicity analysis and prediction "expert system." In previous analyses, the cat-SAR program was able to achieve an overall concordance between observed and predicted values of 92% for a set of chemicals assessed for their ability to induce respiratory hypersensitivity [16], 80%–90% for a set of environmental estrogen mimics [17], and 78%–84% for a set of rat mammary carcinogens [15].

Cat-SAR models are built through a comparison of structural features found amongst categorized compounds in the model's learning set. Generically, these categories are toxicologically active and inactive compounds. Essentially, the cat-SAR approach is transparent in the development of the learning set, the identification of fragments, and the determination of significant or important ones. Moreover, the approach allows user intervention and model optimization throughout the modeling process. This method includes the ability to examine the entire fragment base and to explore and optimize the fragments that have perceived biological relevance.

Moreover, since cat-SAR analyzes categorical data and 2-dimensional fragments rather than intact chemicals, the program can examine noncongeneric datasets that are divided into categories of activity rather than degrees of potency as in the case of quantitative SAR (QSAR). Thus, unlike Hansch and conformational molecular field analysis (CoMFA) approaches that require continuous-type data, cat-SAR works by identifying molecular attributes associated with biological activity by comparing attributes of active (e.g., teratogenic) to inactive (e.g., nonteratogenic) compounds. The models and subsequent predictions based on this dichotomy can then be used to examine structural features associated with teratogenicity and predict the likelihood of teratogenic activity of unknown compounds, respectively.

Overall, the cat-SAR models discussed herein for developmental toxicity demonstrate a high degree of predictivity and mechanistically interpretability and can be useful for screening new drug candidates for developmental toxicity as well as for investigating their therapeutic and toxic mechanisms of action.

2.2.1. Learning Set Development.
The cat-SAR models are built through a comparison of structural features found amongst two designated categories of compounds in the model's learning set. As mentioned, for these analyses the categories were developmental-toxicity and nondevelopmental-toxicity. The cat-SAR learning set consists of the chemical name, its structure as a MOL2 file, and its categorical designation (e.g., one or zero for active and inactive). Typically, organic salts are included as the freebase and simple mixtures and technical grade preparations may be included as the major or active component, metals, metalo-organic compounds and polymers, and mixtures of unknown composition are not included.

2.2.2. In Silico Chemical Fragmentation and the Compound-Fragment Data Matrix.
Using the Tripos Sybyl HQSAR

module, each chemical was fragmented *in silico* into all possible fragments meeting user-specified criteria. HQSAR allows the user to select attributes for fragment determination including atom counts (i.e., the size of the fragments), bond types, atomic connections (i.e., the arrangement of atoms in the fragment), explicit hydrogen atoms, chirality, and hydrogen bond donor and acceptor groups. Fragments can be linear, branched, or cyclic moieties. For analysis of the developmental toxicity dataset the models developed contained fragments between three and seven atoms in size and considered atoms, bond types, and atomic connections.

Upon completion of the fragmentation routine a Sybyl HQSAR add-on is used to produce a compound-fragment data matrix. In the matrix, the rows are intact chemicals and columns are the molecular fragments. Thus for each chemical, a tabulation of all its fragments is recorded across the table rows and for each fragment all chemicals that contain it are tabulated down the columns.

The HQSAR module is not used for statistical analysis or model development. Rather, the compound-fragment matrix is subsequently analyzed with the cat-SAR expert system in order to identify structural features associated with the categorized active and inactive compounds. The cat-SAR program, human developmental toxicity database, and the compound-fragments matrix are available through the corresponding author.

2.2.3. Identifying "Important" Fragments of Activity and Inactivity.

A measure of each fragment's association with biological activity is next determined. To ascertain an association between chemical descriptors (i.e., fragments) and a chemical's activity (or inactivity), a set of rules is used to choose "important" from "unimportant" descriptors. The first selection rule (the Number Rule) is the number of chemicals identified in the learning set that possesses each particular fragment. The second selection rule (the Proportion Rule) is the proportion of active or inactive chemicals that then possesses the particular fragment. Although previously published cat-SAR models required the user to select specific values for the Number and Proportion Rules, a new routine was implemented here to determine optimal values for the Number and Proportion rules. For this exercise the values for the Number Rule were allowed to range between one and eight and the initial values for the Proportion Rule were allowed to range between 0.50 and 0.95.

2.2.4. Predicting Activity.

The resulting list of important fragments can then be used for mechanistic analysis, or to predict the activity of an unknown compound. In the latter circumstance, the model determines which, if any, fragments from the model's learning set the test compound contains. If none are present, no prediction of activity is made for the compound (i.e., no default prediction). If one or more fragments are present, the number of active and inactive compounds containing each fragment is determined. Here the fragment sum (FragSum) method calculates the average probability of the active and inactive fragments contained in a compound and is weighted to the number of active and inactive compounds that go into deriving each fragment. For example, if a compound contains two fragments, one being found in 9/10 active compounds in the learning set (90% active) and the other being found in 3/3 inactive compounds (0% active), the unknown compound will be predicted to be have a probability of activity of 69% (9/10 actives + 0/3 actives = 9/13 actives or 69% chance of activity).

As described, a Cat-SAR prediction of activity or inactivity is based on two separate fragment sets (i.e., the active fragments and the inactive ones) and the predicted activity of a chemical is based on the average probability of all the active and inactive compounds contributing to its structure. Therefore, to classify compounds back to an active or inactive category (i.e., rather than a probability of activity), the program identifies an optimal cutoff point that best separates the prediction of active and inactive compounds based on the probabilistic values of activity derived from a model validation analysis [18]. Depending on the application of the model, the cutoff point that separates active from inactive categorization, for example can be adjusted wherein a model with the best overall concordance can be selected (i.e., a most predictive model), one with equal sensitivity and specificity (i.e., a balanced model that does not overly predictive active compounds at the cost of wrongly predicting inactive one and viceversa), or one with high sensitivity (i.e., a risk averse model).

2.3. Model Validation.

A self-fit, leave-one-out (LOO), and multiple leave-many-out (LMO) cross-validations were conducted for each model (see Table 1). For the self-fit, a model was developed from the complete learning set of 323 compounds and that model was used to predict the activity of each compound in the learning set to determine the general robustness of the model. For the LOO validation, each chemical, one at a time, was removed from the model's total fragment set, and an n-1 model was derived. The activity of the removed chemical was then predicted using the n-1 model. Predicted *vs.* experimental values for each chemical were then compared and the model's overall concordance, sensitivity, and specificity were determined, where,

$$\text{Concordance} = \frac{\text{Correct predictions}}{\text{Total predictions}},$$

$$\text{Sensitivity} = \frac{\text{Correct positive predictions}}{\text{Total positive predictions}}, \quad (1)$$

$$\text{Specificity} = \frac{\text{Correct negative predictions}}{\text{Total negative predictions}}.$$

For the LMO validation, randomly selected sets of 5% of the chemicals were removed from the model, and the n-5% model was derived for each set. The activity of each of the removed chemicals was then predicted using the n-5% model. Predicted versus experimental values for the chemicals in the left out sets were then compared and the n-5% model's concordance, sensitivity, and specificity were determined. This was repeated 10 000 times to compute the model's average concordance, sensitivity, and specificity.

TABLE 1: Fragment summary, self-fit, and cross validation results for the developmental toxicity SAR models.

	Model 1*	Model 2**
Fragments		
Total	26401	26401
Model	4522	3440
Active	1815	1027
Inactive	2707	2413
Self-fit		
Sensitivity	94% (107/114)	98% (94/96)
Specificity	94% (176/177)	100% (156/156)
Concordance	94% (274/291)	99% (250/252)
Leave-one-out		
Sensitivity	78% (90/115)	75% (78/104)
Specificity	79% (141/179)	85% (137/162)
Concordance	79% (231/294)	81% (215/266)
Leave-many-out		
Sensitivity	78% (4.3/5.5)	74% (3.7/5.0)
Specificity	79% (6.8/8.7)	84% (6.6/7.9)
Concordance	78% (11.0/14.2)	80% (10.3/12.9)

*Model 1: Balanced sensitivity and specificity.
**Model 2: Best Concordance.

3. Results and Discussion

3.1. Model Analysis and Validation. Together, two cat-SAR developmental toxicity models were produced. Model 1 was designed to have near-equal sensitivity and specificity while Model 2 was designed to demonstrate the best overall concordance between observed and predicted values (Table 1). Overall, 26 401 unique chemical fragments between 3–7 nonhydrogen atoms were derived for the developmental toxicity database (Table 1). Of these, for Model 1, 1815 were associated with developmental toxicants and 2707 with nontoxicants (4522 total significant fragments). And for Model 2, 1027 were associated with developmental toxicants and 2413 with nontoxicants (3440 total significant fragments).

The self-fit analysis yielded concordance between observed and predicted results of 94% (274/291) for the balanced sensitivity and specificity model (Model 1) and 99% (250/252) for the best concordance model (Model 2). These high concordance rates indicate a robust model wherein there is sufficient structural information contained in the learning set to distinguish between active and inactive compounds. The LOO cross-validation yielded a concordance of 79% for the balanced sensitivity

and specificity model (Model 1) and 81% for the best concordance one (Model 2). The LMO cross-validation yielded an average concordance of 78% for the balanced sensitivity and specificity model (Model 1) and 80% for the best concordance one (Model 2). Overall, since the LOO and LMO validation results are in near agreement, it is estimated that the cat-SAR developmental toxicity model is ~80%, accurate for predicting the developmental toxicity potential of chemicals not included in the model's learning set.

In order to better judge how well these two models performed in general, one can consider the "accuracy" or reproducibility of a standard in vitro toxicological test. For instance, the US National Toxicology Program's (NTP) *Salmonella* mutagenicity database, which is derived from a standardized protocol, has been estimated to be about 85% reproducible (49). Based on these findings wherein the concordance between observed and predicted values for human developmental toxicity is ~80% the cat-SAR human developmental toxicity models thus appear to be as predictive as data used to develop SAR models from standardized in vitro assays.

3.2. Analysis of Propylthiouracil, Methimazole, and Carbimazole. For these analyses, the balanced sensitivity and specificity model (Model 1) was used for analysis of PTU, MMI, and CMI. This model identified 13 fragments within the three antithyroid medications relating to developmental toxicity (Figure 1). Nine of the 13 fragments were considered to be associated with developmental toxicants and 4 of the 13 fragments were considered to be associated with non-developmental toxicants.

Fragments 2184, 2182, and 6241 are found in fluorouracil, for which there is a small amount of human data on developmental toxicity. However the compound is known to interfere with DNA synthesis and has been shown to be a developmental toxicant in rat, mouse, rabbit, hamster, guinea pig, and in nonhuman primates ([19], TERIS MicroMedex accessed 06 July 2009).

Fragments 2184, 2182, and 6229 are found in a series of barbiturates used as sedative-hypnotics and anticonvulsants. Literature on the developmental toxicity of the sedative-hypnotics compounds is contradictory [19]. Amobarbital has been associated in several human studies with cardiovascular malformations although no experimental animal studies appear to have been conducted [19]. Phenobarbital use throughout pregnancy, as an anticonvulsant, has been associated with facial and cardiovascular malformations in humans and experimental animals (TERIS MicroMedex accessed 06 July 2009). Mephobarbital is metabolized to phenobarbital and has been associated with increased risk of facial malformations (TERIS MicroMedex accessed 06 July 2009). There is little animal and human data on metharbital, used as an anticonvulsant, or on butalbital used as a sedative hypnotic.

Fragments 6172 and 6206 are found in thioguanine used as an antineoplastic and in treatment of Crohn's disease. Thioguanine has been demonstrated to produce malformations in experimental animals and in humans ([19], TERIS MicroMedex accessed 06 July 2009).
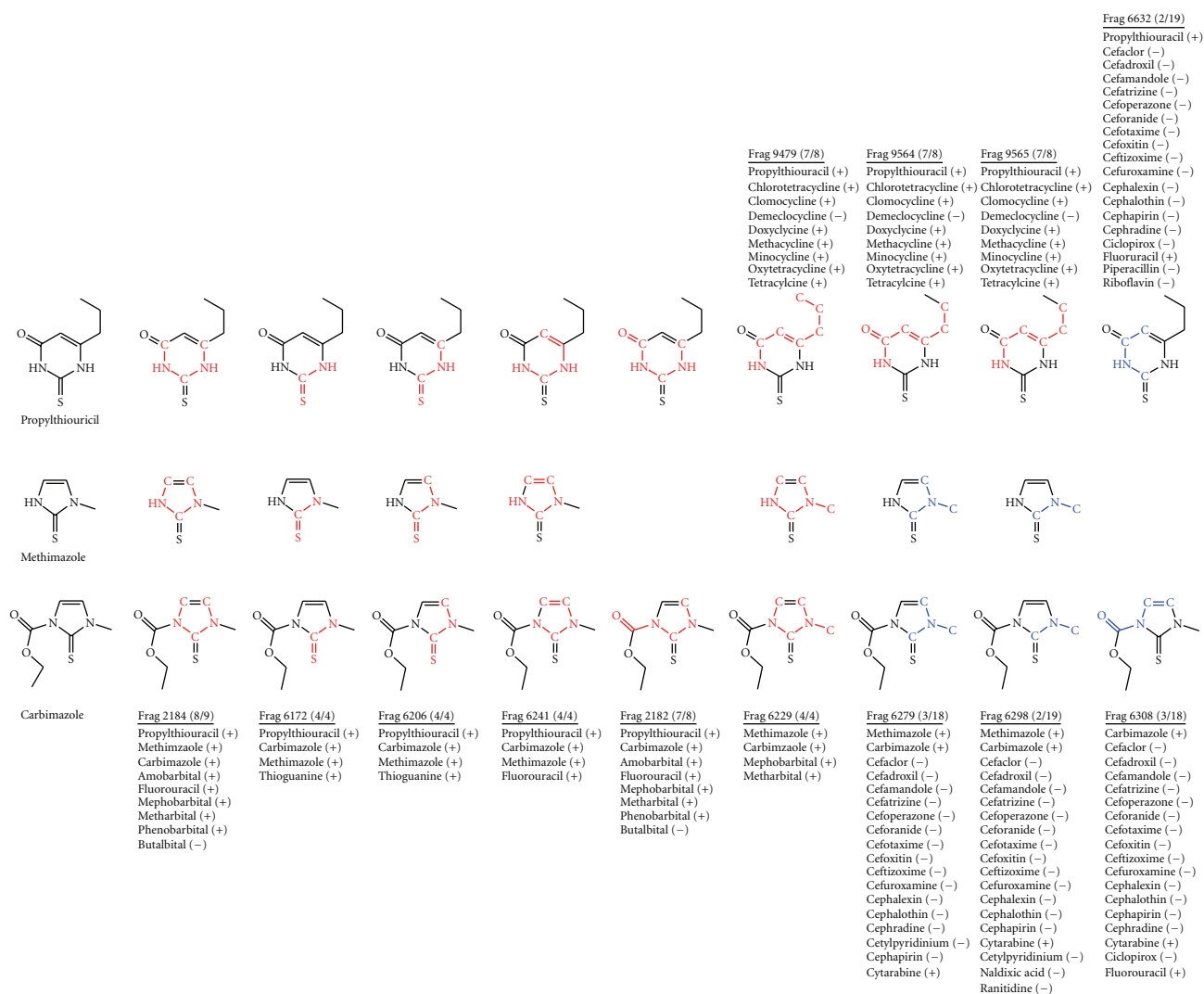
FIGURE 1: Cat-SAR fragments used for developmental toxicity predictions for propylthiouracil, methimazole, and carbimazole. Chemicals in the developmental toxicity database contributing to each fragment are list along with the designation as a developmental toxicant (+) or nontoxicant (−). Red fragments represent developmental toxicity and blue fragments represent nontoxicity.

The additional fragments associated with developmental toxicity, 9479, 9564, and 9565, are found in a series of tetracycline like antibiotics. Tetracycline is identified as a developmental toxicant because of incorporation into teeth and evidence of alteration in bone growth ([19], TERIS MicroMedex accessed 06 July 2009); because of structural similarity and mechanism of action the other tetracyclines are also considered to be developmental toxicants.

There were 4 fragments which were found in molecules thought not to be developmental toxicants: 6279, 6298, 6308, and 6332. The molecules these fragments were found in are cephalosporin, antibiotics and across this class of compounds there is evidence of the absence of risk for developmental toxicity in both experimental animals and humans ([19], TERIS MicroMedex accessed 06 July 2009).

The three antithyroid drugs analyzed represent the limited choices available to the health care provider who is caring for a woman with Graves's disease during pregnancy. These are old drugs, all available as generics; however despite their long history of use there is little data on the developmental risks resulting from use and exposure during pregnancy. All three drugs are capable of crossing the placenta and can cause fetal goiter. Structural malformations are weakly associated with MMI and CMI use during pregnancy. Thus while PTU has been suggested to be without risk of developmental toxicity in humans, that assumption appears to reflect a shared attitude rather than presence of data demonstrating developmental safety. An additional concern with the use of PTU is liver toxicity leading to liver failure and need for transplant [20]. Analysis of data submitted to the US Food and Drug Administration (FDA) suggests that children may be at increased risk for PTU-induced liver failure, and cases have been observed of fetal hepatic toxicity in women treated with PTU (FDA Adverse Events Reporting

System data obtained through the freedom of information act). Additionally there have been cases of infants with malformations observed in women treated with PTU, and structural malformations have been observed in women with Graves's disease who were not treated during pregnancy.

4. Conclusions

The choice of therapeutic strategies for hyperthyroidism during pregnancy is limited. Surgery and radioiodine are typically avoided, leaving propylthiouracil and methimazole in the US [21]. CMI, a metabolic precursor of MMI, is available in some countries outside of the US. In the US PTU is recommended because of concern about developmental toxicity from MMI [22] and CMI [23].

In summary the three drugs available to treat Graves's disease in pregnancy all appear, based on structural analysis, to be capable of producing developmental toxicity. Given the need to treat women with Graves's disease during pregnancy it is essential to develop new molecules with structural attributes which provide suppression of thyroid function while minimizing the risk for developmental toxicity. Based on the results described herein, the cat-SAR method would be a useful approach in screening compounds for developmental toxicity as well as for investigating their therapeutic and potential toxic mechanism of action.

Acknowledgments

References

[1] H. Marx, P. Amin, and J. H. Lazarus, "Hyperthyroidism and pregnancy," *British Medical Journal*, vol. 336, pp. 663–667, 2008.

[2] S. J. Mandel and D. S. Cooper, "The use of antithyroid drugs in pregnancy and lactation," *Journal of Clinical Endocrinology and Metabolism*, vol. 86, no. 6, pp. 2354–2359, 2001.

[3] N. Momotani, J. Y. Noh, N. Ishikawa, and K. Ito, "Effects of propylthiouracil and methimazole on fetal thyroid status in mothers with Graves' hyperthyroidism," *Journal of Clinical Endocrinology and Metabolism*, vol. 82, no. 11, pp. 3633–3636, 1997.

[4] M. L. Browne, S. A. Rasmussen, A. T. Hoyt, et al., "Maternal thyroid disease, thyroid medication use, and selected birth defects in the national birth defects prevention study," *Birth Defects Research A*, vol. 85, no. 7, pp. 621–628, 2009.

[5] N. Momotani, K. Ito, N. Hamada, Y. Ban, Y. Nishikawa, and T. Mimura, "Maternal hyperthyroidism and congenital malformation in the offspring," *Clinical Endocrinology*, vol. 20, no. 6, pp. 695–700, 1984.

[6] G. W. Chan and S. J. Mandel, "Therapy insight: management of Graves' disease during pregnancy," *Nature Clinical Practice Endocrinology & Metabolism*, vol. 3, no. 6, pp. 470–478, 2007.

[7] M. Abalovich, N. Amino, L. A. Barbour, et al., "Management of thyroid dysfunction during pregnancy and postpartum: an Endocrine Society Clinical Practice Guideline," *The Journal of Clinical Endocrinology and Metabolism*, vol. 92, supplement 8, pp. S1–S47, 2007.

[8] O. E. Okosieme, H. Marx, and J. H. Lazarus, "Medical management of thyroid dysfunction in pregnancy and the postpartum," *Expert Opinion on Pharmacotherapy*, vol. 9, no. 13, pp. 2281–2293, 2008.

[9] F. R. Jelovsek, D. R. Mattison, and J. J. Chen, "Predicition of risk for human developmental toxicity: how important are animal studies for hazard identification," *Obstetrics and Gynecology*, vol. 74, pp. 624–636, 1989.

[10] A. F. Olshan, D. R. Mattison, and T. S. Zwanenburg, "International Commission for Protection Against Environmental Mutagens and Carcinogens. Cyclosporine A: review of genotoxicity and potential for adverse human reproductive and developmental effects. Report of a working group on the genotoxicity of cyclosporine A, August 18," *Mutation Research*, vol. 317, no. 2, pp. 163–173, 1994.

[11] N. Takihi, H. S. Rosenkranz, G. Klopman, and D. R. Mattison, "Structural determinants of developmental toxicity," *Risk Analysis*, vol. 14, pp. 649–657, 1994.

[12] M. Ghanooni, D. R. Mattison, Y. P. Zhang, O. T. Macina, H. S. Rosenkranz, and G. Klopman, "Structural determinants associated with risk of human developmental toxicity," *American Journal of Obstetrics and Gynecology*, vol. 176, no. 4, pp. 799–806, 1997.

[13] J. Gómez, O. T. Macina, D. R. Mattison, Y. P. Zhang, G. Klopman, and H. S. Rosenkranz, "Structural determinants of developmental toxicity in hamsters," *Teratology*, vol. 60, no. 4, pp. 190–205, 1999.

[14] F. R. Jelovsek, D. R. Mattison, and J. J. Chen, "Prediction of risk for human developmental toxicity: how important are animal studies for hazard identification?" *Obstetrics and Gynecology*, vol. 74, pp. 624–636, 1989.

[15] A. R. Cunningham, S. T. Moss, S. A. Iype, G. Qian, S. Qamar, and S. L. Cunningham, "Structure-activity relationship analysis of rat mammary carcinogens," *Chemical Research in Toxicology*, vol. 21, no. 10, pp. 1970–1982, 2008.

[16] A. R. Cunningham, S. L. Cunningham, D. M. Consoer, S. T. Moss, and M. H. Karol, "Development of an information-intensive structure-activity relationship model and its application to human respiratory chemical sensitizers," *SAR and QSAR in Environmental Research*, vol. 16, no. 3, pp. 273–285, 2005.

[17] A. R. Cunningham, D. M. Consoer, S. A. Iype, and S. L. Cunningham, "A structure-activity relationship analysis for the identification of environmental estrogens: the categorical-SAR (cat-SAR) approach," in *Endocrine Disruption Modeling*, J. Devillers, Ed., pp. 173–198, CRC Press, Boca Raton, Fla, USA, 2009.

[18] A. R. Cunningham, H. S. Rosenkranz, and G. Klopman, "Identification of structural features and associated mechanisms of action for carcinogens in rats," *Mutation Research*, vol. 405, no. 1, pp. 9–28, 1998.

[19] J. L. Schardein, *Chemically Induced Birth Defects*, Marcel Dekker, New York, NY, USA, 2000.

[20] S. A. Rivkees and D. R. Mattison, "Propylthiouracil (PTU) hepatoxicity in children and recommendations for discontinuation of use," *International Journal of Pediatric Endocrinology*, vol. 2009, Article ID 132041, 8 pages, 2009.

[21] P. Laurberg, C. Bournaud, J. Karmisholt, and J. Orgiazzi, "Management of Graves' hyperthyroidism in pregnancy: focus on both maternal and foetal thyroid function, and caution against surgical thyroidectomy in pregnancy," *European Journal of Endocrinology*, vol. 160, no. 1, pp. 1–8, 2009.

[22] P. Barbero, R. Valdez, H. Rodriguez, et al., "Choanal atresia associated with maternal hyperthyroidism treated with methimazole: a case-control study," *American Journal of Medical Genetics A*, vol. 146, no. 18, pp. 2390–2395, 2008.

[23] L. Kannan, S. Mishra, R. Agarwal, V. Kartikeyan, N. Gupta, and M. Kabra, "Carbimazole embryopathy-bilateral choanal atresia and patent vitello-intestinal duct: a case report and review of literature," *Birth Defects Research A*, vol. 82, no. 9, pp. 649–651, 2008.